A 3D Approach to Infrastructure-Free Localization in Large Scale Warehouse Environments

Dylan Schwesinger and John Spletzer

***Abstract*— In this paper, we present a method for infrastructure-free localization of Automated Guided Vehicles (AGVs) in a warehouse environment. To accomplish this, our approach leverages 3D data for both mapping and feature segmentation. First, a 3D reconstruction of the warehouse is created to extract salient natural features — in this case the shelving uprights — as landmarks. Next, the map-based localization approach leverages 3D LIDAR to enable 3D feature-to-landmark matching which minimizes the potential for data association errors. In our experiments in a representative warehouse environment, we demonstrated a localization accuracy of approximately 2cm without the use of retroreflector targets. Furthermore, 100% of visible landmarks were detected and there were no false positives.**

I. Introduction

In warehouse environments, manual forklifts (aka "lift trucks") are the workhorses of material handling [1]. In 2013 alone it was estimated that orders for over 1 million forklifts were placed [2]. A typical work cycle involves ferrying palletized goods and materials between storage (either racks and shelves or block storage areas) and trucks for receiving or shipment. Despite their flexibility and effectiveness in material handling tasks, they are not without their shortcomings. These include operating efficiency, high energy consumption, and safety considerations. These concerns have led to the rise of Automated Guided Vehicle (AGV) systems for automated pallet transport and storage [3], [4], [5], [6]. While they lack the adaptability of human operators, for repetitive material handling tasks they are far more efficient - often operating around the clock, and with reduced product damage. Also in large part due to rigorous ANSI safety standards regulating their design and use [7], they are far safer in practice.

AGVs are not merely vehicles, but autonomous *systems* in the full sense of the word. The components of an AGV system include one or more automated vehicles, a *localization system* which provides precise position and orientation estimates for the vehicles, a *route map* or network which delineates the guidepaths where the AGVs can travel, and a centralized controller which coordinates between and assigns specific tasks to the individual vehicles [8]. While all of these components are essential for AGV operation, it can be argued that the enabling technology is the localization system as it answers the first fundamental question which an autonomous vehicle must ask (i.e., Where am I?). There are multiple approaches to AGV localization [9]. The first systems relied upon either wire guidance, where wires are embedded in the floor and sensed inductively, or inertial guidance with magnets placed in the floor which act as landmarks to reset the dead reckoning system for overcoming drift. For contemporary AGV installations, the "gold standard" for localization is arguably laser guidance. In this paradigm, each AGV vehicle is fitted with a 2D scanning LIDAR system. As the AGV navigates, the LIDAR tracks precisely placed retro-reflector targets in the environment which serve as landmarks. Retro-reflectors are used so that the target's reflectivity can be used as a filter for robust landmark segmentation. Otherwise, attempting to segment targets in a 3D world using a 2D sensor is extremely error prone. In practice the approach works quite well, providing sub-centimeter positioning accuracy.

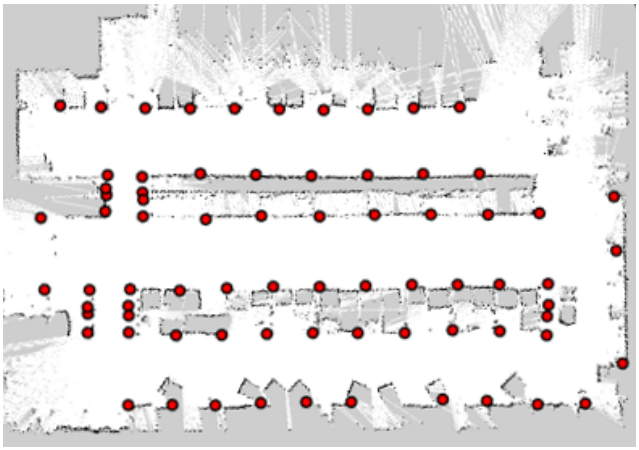

Fig. 1. An occupancy grid map of an approximately 20×35 meter section of warehouse. The red circles denote a feature map composed of the vertical pallet rack supports.

A significant drawback of each of the aforementioned approaches is that they require modifications to the warehouse infrastructure. In the case of laser-guidance, hundreds or even thousands of targets need to be accurately placed within the environment. The large number is necessary to ensure the required localization accuracy can be achieved from any location in the facility, as industrial environments are often very cluttered [8]. The result is not only significant per-vehicle costs, but also very high installation costs. Furthermore, the approach has limited flexibility in that changes to a facility's layout may necessitate significant modifications to the reflector installation.

It should then come as no surprise that there is significant interest in *infrastructure-free* solutions to AGV localization.

The authors are with the VADER Laboratory in the Computer Science and Engineering Department of Lehigh University, Bethlehem, Pennsylvania, United States dts211@lehigh.edu, spletzer@cse.lehigh.edu

Indeed, in their roadmap for boosting the use of AGVs in industrial applications, the authors in [8] specifically cite localization based upon natural landmarks as a key enabling technology. Both academic and industry researchers have taken notice. Vision-based solutions have been proposed in [9], [10], and most significantly by Seegrid Corporation [6]. These approaches come with the pros (e.g., lower-cost, passive sensors) and cons (e.g., limited robustness to illumination changes, inability to operate in darkness) associated with vision sensors. More recently, researchers have begun investigating the use of LIDAR systems to localize in large-scale warehouse environments by leveraging maps generated *a priori* [11], [12]. However, the authors' use of a 2D LIDAR for localization in a cluttered industrial environment necessitated a contour-based approach. Of concern is that palletized goods and materials may make up a significant portion of the LIDAR scan (contour), and the robustness of the approach to permutations in pallet placement and density has only seen limited validation.

With these insights, we argue that:

1) Robust, infrastructure-free localization in a cluttered industrial environment motivates 3D LIDAR systems for perception
2) Feature based localization approaches are preferable, as they are more stable and less dependent upon variability in warehouse inventory, and
3) Using 3D LIDAR sensors, there are sufficient natural 3D features in warehouse environments that the use of retro-reflector targets can be eliminated.

Unfortunately, the accessibility of such an approach has been constrained due to the limited availability of suitable sensing technology. However this is rapidly changing, and AGV localization now appears ready for a paradigm shift.

Based on our previous work in landmark map-based localization in large scale urban environments [13], we propose an adaptation of the approach for large scale indoor warehouse environments. In urban environments, we used ubiquitous pole-like features, such as lamp posts, parking meters, and decorative trees, as natural landmarks in our map. Similarly, in warehouse environments, we propose using the pole-like pallet rack supports as landmarks. In addition to the new application area, the main extensions of this work include the integration of the Velodyne VLP-16 LIDAR, a comparison of alternate mapping paradigms, and methods used to segment the pallet rack supports from 3D representations of the warehouse environment.

II. Related Work

Mapping and localization in static environments is a mature area in robotics with a rich body of literature. As such, we limit the scope of discussion to the state-of-the-art in the specific context of warehouse environments.

In [14], the authors propose an infrastructure-free framework for warehouse navigation that uses a topological map rather than a globally consistent metric map. Their approach is to use a monocular camera to track the texture on the floor and the map is a locally consistent pose graph representation where each pose in the graph has an associated image. Localization is performed by matching the current camera frame to a likely subset of the pose graph. Due to the differences in map representation, our approach is not directly comparable.

More similar to our work in terms of map representation is a body of research [11], [12], [15] under the PAN-Robots project [16] funded by the European Union. In [12] the authors use the GMapping algorithm [17] to build an 2D occupancy grid based map representation. We similarly use the GMapping algorithm, but only as a means to build a globally consistent 3D reconstruction of the environment. In [15], the authors extend the work to include GraphSLAM [18] based mapping approach where existing retro-reflectors were used as landmarks. This is similar to our approach in that the map representation is feature-based, however our landmarks are based on natural 3D features of the environment, rather than artificial 2D features. While a stated goal in [12] is to enable 3D mapping, all the mapping and localization work is 2D based. There is only a brief mention that 3D mapping is possible and is left as future work.

In [12], [15] they use a contour-based, adaptive Monte Carlo localization (AMCL) approach [18] where they use P-L-ICP [19] for visual odometry. In contrast, our localization approach focuses on naturally occurring warehouse features. By leveraging th latest in 3D LIDAR systems for localization, specifically the Velodyne VLP-16 LIDAR [20], robust, real-time 3D feature segmentation and landmark association is achieved. This 3D-to-3D feature mapping provides a significant advantage over 2D contour approaches in terms of system robustness. Furthermore, since the features we are tracking are temporally invariant for a given warehouse layout, we expect this approach will provide more stable and consistent localization performance.

III. Map Generation

The map representation used by the system is feature-based. In this work, we chose to use the vertical, pole-like supports of pallet racks as landmark features, which we will refer to simply as pole features. The mapping process consists of three main steps: (1) collect dense laser scans of the environment, (2) register the laser data to a common coordinate frame to create a 3D point cloud reconstruction of the environment, and (3) segment salient 3D features from the registered point cloud to create a landmark map.

A. The Mapping Trike Platform

To acquire dense laser scans of the environment, we utilized our Mapping Trike platform [13]. Details of the sensor suite are repeated here for convenience. Pose estimation was provided by a Microstrain 3DM-GX3-45 inertial measurement unit (IMU) in conjunction with two 4096 cycles per revolution (CPR) resolution encoders mounted on the rear wheels. Two SICK LMS291-S14 LIDARs were mounted facing each side for the purpose of creating the 3D reconstruction. A single SICK LMS291-S05 LIDAR was mounted parallel to the ground plane facing the rear for the purpose of global pose corrections.

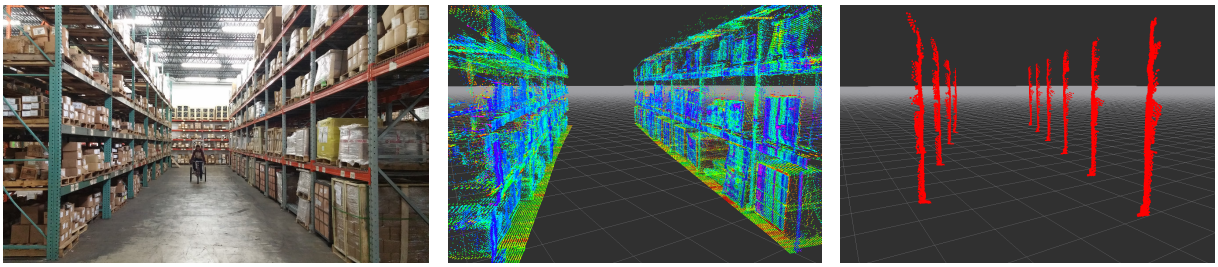

Fig. 2. (Left) One aisle of the warehouse. (Center) A point cloud reconstruction. (Right) Landmark segmentation for feature map generation.

B. Coordinate Frame Registration

We note that although we are creating a 3D reconstruction of the warehouse environment and tracking 3D features, AGV localization will be on the plane. As a result, the mapping goal is register the 2D positions of the landmarks to a common global coordinate frame. There are many mature solutions to the Simultaneous Localization and Mapping (SLAM) problem for learning 2D maps. In this work, we chose to use the GMapping algorithm [17]. The main reason for this decision was convenience as a high quality open source solution implementation is available [21]. However, any SLAM algorithm capable of recovering the trike's trajectory could be a suitable replacement.

The GMapping algorithm is a particle filter based approach to learn a 2D occupancy grid representation of the environment from horizontal planar laser scans and odometry measurements. Laser scans from the rear-facing LIDAR were used as input. For the odometry measurements, an extended Kalman filter (EKF) was used which incorporated the vehicle kinematics in the predictive step and data from the encoder and IMU for the corrective step. Since the floor of a warehouse environment is roughly planar, we used a 3 degree of freedom motion model of the form:

$$\begin{pmatrix} x \\ y \\ \theta \end{pmatrix}_{k+1} = \begin{pmatrix} x \\ y \\ \theta \end{pmatrix}_{k} + \begin{pmatrix} \cos\theta & 0 \\ \sin\theta & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta_l \\ \Delta_a \end{pmatrix},$$

where x and y are Cartesian coordinates, θ is the yaw angle, Δ_l and Δ_a were the linear and angular displacements over some small discrete time step denoted as k. Note that there is an unmodeled kinematic constraint from the trike's front wheel, but this was not needed for mapping purposes.

The GMapping algorithm learns a 2D occupancy grid representation of the environment. However, our goal is to learn a 2D feature-based representation by creating a 3D reconstruction of the environment, segmenting the landmarks from this representation, and then registering the landmarks to the 2D plane. We used a structure-from-motion based approach to register the side-facing LIDAR scans to the GMapping result, which only required logging the pose corrections at each time step of the GMapping algorithm.

C. Landmark Segmentation

With the corrected trike poses, a globally consistent 3D point cloud reconstruction of the environment could be constructed as depicted in Fig. 2 (center). The goal was to segment the pole features from this reconstruction. First each 2D laser scan was transformed to a coordinate frame where the x axis was parallel to the ground plane, here denoted as a sequence of consecutive points, $S = (p_1, \ldots, p_k)$, where k is the number of points in a scan and each p_i is a vector of coordinates, $[x, y]^T$.

Then each scan was clustered into subsets where the points in a valid cluster satisfied three conditions:

1) $\|p_{i+1} - p_i\| < \alpha$, where α is the distance tolerance between points,
2) $\|a^T p_i - f(S)\| < \beta$, where $a = [0, 1]^T$, $f(S)$ is the median y value of the points in scan S and β is a distance tolerance to $f(S)$,
3) $\|p_{n_c} - p_{1_c}\| > \gamma$, where p_{1_c} and p_{n_c} are the first and last points in cluster c respectively, and γ is a tolerance on the length of a cluster.

The assumption was that a pallet rack support was the tallest contiguous vertical object in a scan that contained one. Hence, the first condition validates "contiguousness", the second condition validates "verticality", since the median value should lie on the pole feature, and the third condition validates "tallness". In this work, the α and β values were both set to 2 cm, and γ was set to 30 cm. These values were chosen based on the LIDAR's anglular resolution and mounted orientation.

The points from all valid clusters from all laser scans were transformed to the 3D global coordinate frame based on the corrected trike poses and combined into a set M. The points in M were clustered into subsets that satisfied the condition $\|p_i - p_j\| < \epsilon$ where ϵ is the maximal intra-cluster distance between points. In this work, ϵ was set to 5 cm based on the trike driving approximately 1 m/s during mapping. Any cluster that had a bounding box height greater than 2 meters was considered to be a landmark. This value was chosen based on the LIDAR's field of view and the expected minimum range to a landmark. A depiction of accepted landmark clusters is shown in Fig. 2 (right). The 2D landmark map was then synthesized by projecting centroids of each accepted cluster to the ground plane.

IV. Map-based Localization

This section describes the proposed AGV localization procedure, which assumes that a landmark map is available. The main components of the system are perception and

localization. The perception component handles landmark segmentation and the localization component maintains an estimate of the AGV's pose within the landmark map. A description of each of these components and the development platform now follows.

A. The Development Platform

Since an AGV was not initially available to us, to demonstrate feasibility we employed our Smart Wheelchair (SWS) platform as a surrogate [13]. The SWS was equipped with a mast-mounted Velodyne VLP-16 for exteroceptive sensing. The VLP-16 was chosen as it provides accurate, real-time 360° 3D measurements. It features 16 laser sensors with a vertical field of view of 30°. We also note that the VLP-16 costs less ($\approx$ \$8000 USD) than single-beam LIDARs currently used on AGV reflector-based localization systems. The SWS also provided odometry information from wheel encoders and IMU measurements.

B. Landmark Segmentation

We denote a 360° VLP-16 laser scan as a matrix of the form:

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,k} \\ \vdots & \vdots & \ldots & \vdots \\ p_{16,1} & p_{16,2} & \cdots & p_{16,k} \end{pmatrix},$$

where k is the number of scans, $p_{i,j}$ is a 3D coordinate. Each row corresponds to a laser scan and each column corresponds to an azimuth angle.

The first step is to correct the measurements in P to compensate for vehicular motion. We assume constant motion between odometry measurements and a constant rotational velocity of the sensor. For each column j in P, the corrected translation is given by $lerp(w_0, w_1, \eta)$ where $lerp(\cdot)$ is the linear interpolation operation, w_0 and w_1 are the measured vehicle position at the beginning and end of the scan respectively, and $\eta = (j-1)/(k-1)$ is the interpolation parameter. Similarly, the corrected rotation for each column j in P is given by $slerp(q_0, q_1, \eta)$, where $slerp(\cdot)$ is the spherical linear interpolation operation and q_0 and q_1 are quaternion representations of the measured vehicle yaw at the beginning and end of the scan respectively.

Next, the points in each column j in P are associated to 2D polar coordinates, $(\rho_{i,j}, \phi_{i,j})$, based on the x and y components of each point, assuming that the $x-y$ plane is the ground. Then a set of valid points is created:

$$V = \{p_{i,j} \mid \|\rho_{1,j} - h(j)\| < \delta \wedge \ldots \wedge \|\rho_{16,j} - h(j)\| < \delta\},$$

where $h(\cdot)$ is the median ρ value in column j and δ is a distance tolerance from this value. In this work, the value of δ was set to 3 cm based on the accuracy of the VLP-16. The assumption was that due to the limited vertical field of view and *a priori* knowledge of the sensor position that all 16 scans at a given azimuth would hit a pole feature.

The points in V were then clustered into subsets that satisfied the condition $\|\pi_{xy}(p_i) - \pi_{xy}(p_j)\| < \xi$ where $\pi_{xu}(p)$ is the $x-y$ projection of point p and ξ is the maximal intra-cluster distance between points. In this work, ξ was set to 3 cm based on the accuracy of the VLP-16. Each cluster was then validated based on the geometry of its oriented bounding box. Clusters were rejected if: the height was less than 1.2 meters, the max(width, depth) was greater than 0.5 meters, or the ratio of the height to the max(width, depth) was less than 5. These values were chosen based on vertical resolution of the VLP-16, the expected minimum distance to detect a landmark, and the expected width of a landmark. Fig. 3 depicts the resulting segmentation. The segmentation operated in real-time with a VLP-16 scan rate of 5 Hz.

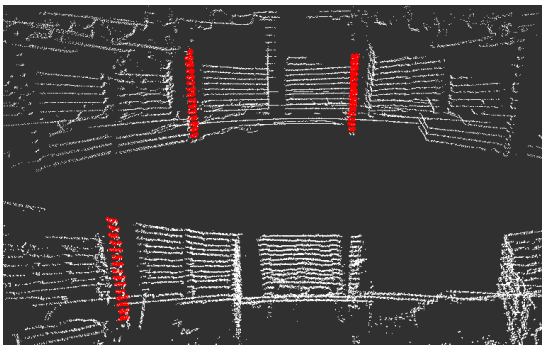

Fig. 3. Real-time landmark segmentation for localization using the Velodyne VLP-16. Detected landmarks are highlighted in red.

C. Localization Procedure

Localization of the SWS was performed using the same particle filter based approach described in our previous work [13]. In short, we implemented a variant of the feature-based FastSLAM 2.0 algorithm [22] that performed localization, but no mapping. We found that due to the improved proposal distribution, fewer particles were needed to maintain localization. In this work, we used 20 particles for our localization experiments. This compares favorably to the AMCL approach in [12] which required hundreds of particles for effective localizaion. For pose estimation the mean over the particle set was used.

V. Experiments

To demonstrate the effectiveness of the proposed method in its intended environment, a map was constructed of an approximately 20 $\times$ 35 meter section of warehouse shown in Fig. 1. The following sections describe the mapping results and the client localization results. We leveraged the Robot Operating System (ROS) [23] framework for our implementation and also utilized the Point Cloud Library (PCL) [24] for processing point cloud data from the exteroceptive sensors.

A. Mapping

To construct the map, data were collected by driving the Mapping Trike through the warehouse at a rate of approximately 1 m/s. In conjunction with the LMS291 LIDAR scan rate of 75 Hz, this gave us a vertical scan for each 1-2 centimeters traveled which was sufficiently dense

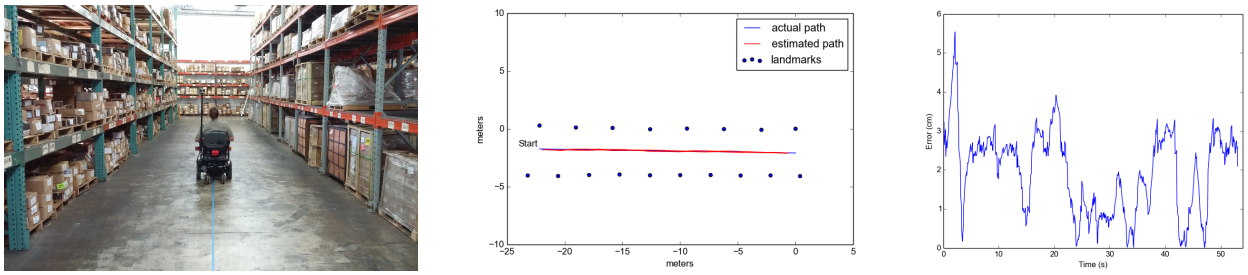

Fig. 4. (Left) The localization accuracy experiment set up with the SWS equipped with a Velodyne VLP16. (Center) The actual vs. estimated path. (Right) Absolute error vs. time. The mean absolute error was 1.9 cm.

for our landmark segmentation procedure. An advantage of the Mapping Trike platform in this scenario was that the maneuverability allowed us to collect data without disrupting normal warehouse operations.

To validate the effectiveness of the landmark segmentation of the Mapping Trike, the number of landmarks was counted by hand as a ground truth measure. This value was compared against the map generated by the mapping process. In total, 74 of 74 visible landmarks were successfully segmented. Furthermore, no false positives were detected. From this, we report that in our test environment the proposed mapping approach had a 100% true positive rate for segmenting landmarks, and a 0% false positive rate. While there were no ground truth measurements of the true landmark coordinates to determine a quantitative measure of accuracy, the generated map was qualitatively consistent to the environment based on the spacing between landmarks and the ability of the SWS to localize.

A second mapping experiment was performed to evaluate the viability of mapping with the SWS platform. The motivation for this was to investigate the potential for eliminating the special purpose mapping vehicle. In other words, could an AGV equipped with a VLP-16 LIDAR create its own map? While the accuracy of the VLP-16 in conjunction with the motion estimation error of the platform was initially deemed unsuitable to create an accurate 3D reconstruction, the SWS (our AGV surrogate) had the advantage of being able to detect each landmark multiple times from different poses. Thus, a given landmark position estimate could be derived from multiple constraints.

To build the map, data were logged while driving several loops through the warehouse (the same data set used in Fig. 5), landmarks were segmented using the approach in Section IV-B, and then a classic EKF SLAM formulation [25] was used to learn the map. The SWS generated map contained all the ground truth landmarks with no false positives.

To quantitatively compare the maps a minimization of the form:

$$\underset{R,t}{\operatorname{argmin}} f(R,t) = \frac{1}{N}\sum_{i=1}^{N} \|x_i - Ry_i - t\|^2,$$

where x_i and y_i are the corresponding points, N is the number of points, R is a rotation matrix in $SO(2)$, and t is a translation vector, was solved to obtain the obtain the R and t to align the points. After the points were aligned the distances between corresponding points were computed and used as a quantitative measure of map similarity. This resulted in a mean distance of 4.3 cm with a standard deviation of 3.5 cm. Again, while we lack absolute ground truth, the relative consistency of the two maps indicates the potential for an AGV equipped with a Velodyne VLP-16 LIDAR to create its own landmark map.

B. Localization Accuracy

In an attempt to quantify localization accuracy, we performed an experiment where we marked a line down the center of the middle aisle with blue tape, shown in Fig. 4 (left). The line served as "ground truth" and was measured by hand relative to the landmark positions. The operator then attempted to drive the SWS straight down the line, turn in place at the end, and drive straight back to the starting position. Results of localization are shown in Fig. 4. The center figure depicts the path traveled and the right figure depicts the absolute error to the target line over time. The average path error was 1.9 cm with a standard deviation of 1.1 cm and a median error of 2.3 cm. While this is in excess of our target centimeter-level accuracy, we note that the analysis assumes that the line was perfectly straight (it was not) and the SWS operator drove down the center of the line (he did not). This is evidenced in Fig. 4 (right), where there is a noticeable symmetry about the 30 second mark when the turn in place maneuver was performed. This is likely due to a bias error in the placement of the tape line. As a result, centimeter-level localization accuracy may have been achieved. We acknowledge that this was a limited test, but argue that it demonstrates the potential of our approach to achieve the necessary localization accuracy for an infrastructure-free warehouse environment.

To investigate the ability to maintain localization over time, a second experiment was performed where the SWS was driven around several arbitrary loops in the map at an average speed of 1.0 m/s for approximately six minutes. The estimated path is shown in Fig. 5. The starting pose was approximately [-19 -9 0]^T, and the total distance traveled was 353 meters. There were an average 9.5 landmark observations per meter traveled. Every visible landmark was detected, and there were no false positive (a feature being incorrectly associated with a landmark) as accurate localization

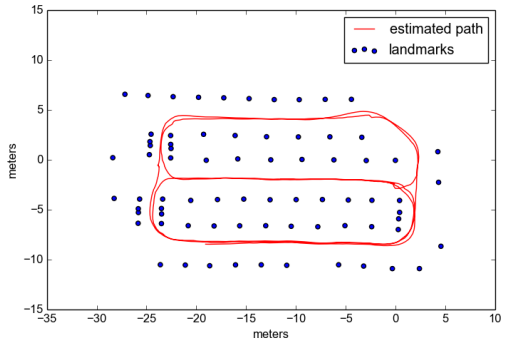

Fig. 5. The estimated real-time corrected path where the SWS was driven around several arbitrary loops within the warehouse map.

was maintained throughout the trial. A video showing the 3D reconstruction and landmark segmentation for part of this experiment can be viewed at `https://www.youtube.com/watch?v=B73tgGoT6Ms&feature=youtu.be`.

VI. Conclusions & Future Work

In this work, we presented a 3D map-based approach to infrastructure-free localization of an AGV in a warehouse environment. Preliminary results support our hypothesis that there are sufficient natural 3D landmarks to support a robust, feature-based localization approach. During warehouse mapping, visible pole features were segmented with 100% reliability, and with a 0% false positive rate. During localization trials, all visible landmarks were detected, and again no false positive landmarks were identified. The demonstrated accuracy of the localization system was approximately 2cm.

This research was inspired in part by our work in outdoor localization and mapping [13]. One significant improvement that we observed herein was the potential for eliminating the special purpose mapping vehicle. Compared to the actuated Hokuyo UTM-30LX used in our previous work, the VLP-16 has over 3X the range, 2X the number of effective beams, up to 2.5X the angular resolution, and greater fields-of-view. Our experimental results indicate that these improvements are sufficient to enable mapping with the AGV platform itself, which dramatically simplifies the logistics for a real-world implementation.

We are currently working with an AGV company to benchmark our approach against a traditional LIDAR/retro-reflector localization system in a representative warehouse environment.

Acknowledgment

The authors would like to thank Tom Panzarella from Love Park Robotics and Armon Shariati from the GRASP Laboratory at the University of Pennsylvania for their help in this work.

References

[1] "Forklift Trucks - the Backbone of the Industry," *MHEDA Journal Online*, 2004.

[2] Modern Material Handling. (2014, August) Top 20 Lift Truck Suppliers, 2014. [Online]. Available: http://www.mmh.com/article/top_20_lift_truck_suppliers_2014

[3] JBT Automated Guided Vehicles. Pallet Handling Automatic Guided Vehicles. [Online]. Available: http://www.jbtc-agv.com/en/Solutions/Applications/Pallet-Handling

[4] Dematic, "Automated Pallet Transport and/or Storage (AGV)," www.dematic.com/en-US/Supply-Chain-Solutions/By-Vertical-Market/Typical-Solutions/Automate-Pallet-Transport-and-Storage-AGV.

[5] E. Automation, "Egemin Automation - Automated Guided Vehicles and Integrated Material Handling Solutinos," http://www.egeminusa.com/.

[6] Seegrid, "Seegrid Vision - the Power to Transform," http://www.seegrid.com/.

[7] ANSI, *ANSI/ITSDF B56.5-201 2Safety standard for driverless, automatic guided industrial vehicles and automated functions of manned industrial vehicles*. American National Standards Institute, 2012.

[8] L. Sabattini, V. Digani, C. Secchi, G. Cotena, D. Ronzoni, M. Foppoli, and F. Oleari, "Technological roadmap to boost the introduction of agvs in industrial applications," in *IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*, 2013, pp. 203–208.

[9] A. Kelly, B. Nagy, D. Stager, and R. Unnikrishnan, "An Infrastructure-Free Automated Guided Vehicle Based on Computer Vision," *IEEE Robotics and Automation Magazine*, 2007.

[10] L. Sabattini, A. Levratti, F. Venturi, E. Amplo, C. Fantuzzi, and C. Secchi, "Experimental comparison of 3D vision sensors for mobile robot localization for industrial application: Stereo-camera and RGB-D sensor," in *12th International Conference on Control Automation Robotics & Vision, ICARCV 2012, Guangzhou, China*, December 2012, pp. 823–828.

[11] C. Reinke and P. Beinschob, "Strategies for Contour-Based Self-Localization in Large-Scale Modern Warehouses," in *IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*, September 2013, pp. 223–227.

[12] P. Beinschob and C. Reinke, "Advances in 3d data acquisition, mapping and localization in modern large-scale warehouses," in *Intelligent Computer Communication and Processing (ICCP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 265–271.

[13] D. Schwesinger, A. Shariati, C. Montella, and J. Spletzer, "A smart wheelchair ecosystem for autonomous navigation in urban environments," *Autonomous Robots*, pp. 1–20, 2016. [Online]. Available: http://dx.doi.org/10.1007/s10514-016-9549-1

[14] M. Gadd and P. Newman, "A framework for infrastructure-free warehouse navigation," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015, pp. 3271–3278.

[15] P. Beinschob and C. Reinke, "Graph slam based mapping for agv localization in large-scale warehouses," in *Intelligent Computer Communication and Processing (ICCP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 245–248.

[16] PAN-Robots. [Online]. Available: http://www.pan-robots.eu

[17] G. Grisetti, C. Stachniss, and W. Burgard, "Improved techniques for grid mapping with rao-blackwellized particle filters," *Robotics, IEEE Transactions on*, vol. 23, no. 1, pp. 34–46, 2007.

[18] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. MIT press, 2005.

[19] A. Censi, "An icp variant using a point-to-line metric," in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*. IEEE, 2008, pp. 19–25.

[20] Velodyne. VLP-16 LiDAR. [Online]. Available: http://velodynelidar.com/vlp-16.html

[21] G. Grisetti, C. Stachniss, and W. Burgard. GMapping. [Online]. Available: https://www.openslam.org/gmapping.html

[22] M. Montemerlo and S. Thrun, "Fastslam 2.0," *FastSLAM: A Scalable Method for the Simultaneous Localization and Mapping Problem in Robotics*, pp. 63–90, 2007.

[23] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "Ros: an open-source robot operating system," in *ICRA workshop on open source software*, vol. 3, no. 3.2, 2009, p. 5.

[24] R. B. Rusu and S. Cousins, "3d is here: Point cloud library (pcl)," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1–4.

[25] R. Smith, M. Self, and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," in *Autonomous robot vehicles*. Springer, 1990, pp. 167–193.